

RECEIVED

JAN 02 2003

TECH CENTER 1600/2900



1600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/846,637C

DATE: 12/26/2002

TIME: 13:32:53

Input Set : A:\2502seq.003

Output Set: N:\CRF4\12262002\I846637C.raw

3 <110> APPLICANT: Jensen, Michael
5 <120> TITLE OF INVENTION: Selection Systems for Genetically
6 Modified Cells
8 <130> FILE REFERENCE: 24751-2502
10 <140> CURRENT APPLICATION NUMBER: US/09/846,637C
11 <141> CURRENT FILING DATE: 2001-04-30
13 <160> NUMBER OF SEQ ID NOS: 40
15 <170> SOFTWARE: FastSEQ for Windows Version 4.0
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 1654
19 <212> TYPE: DNA
20 <213> ORGANISM: Homo sapien
22 <220> FEATURE:
23 <221> NAME/KEY: CDS
24 <222> LOCATION: (48)...(1589)
25 <223> OTHER INFORMATION: Human Wild-type Inosine Monophosphate Dehydrogenase
26 II (IMPDH II)
28 <300> PUBLICATION INFORMATION:
29 <301> AUTHORS: Collart, F.R. and Huberman, E.
30 <302> TITLE: Cloning and sequence analysis of the human and
31 <303> JOURNAL: J. Biol. Chem. (1988)
32 <304> VOLUME: 263
33 <306> PAGES: 15769-15772
35 <400> SEQUENCE: 1
36 gaattcgggc ggtcctcgga gacacgcggc ggtgtcctgt gttggcc atg gcc gac 56
37 Met Ala Asp
38 1
40 tac ctg att agt ggg ggc acg tcc tac gtg cca gac gac gga ctc aca 104
41 Tyr Leu Ile Ser Gly Gly Thr Ser Tyr Val Pro Asp Asp Gly Leu Thr
42 5 10 15
44 gca cag cag ctc ttc aac tgc gga gac ggc ctc acc tac aat gac ttt 152
45 Ala Gln Gln Leu Phe Asn Cys Gly Asp Gly Leu Thr Tyr Asn Asp Phe
46 20 25 30 35
48 ctc att ctc cct ggg tac atc gac ttc act gca gac cag gtg gac ctg 200
49 Leu Ile Leu Pro Gly Tyr Ile Asp Phe Thr Ala Asp Gln Val Asp Leu
50 40 45 50
52 act tct gct ctg acc aag aaa atc act ctt aag acc cca ctg gtt tcc 248
53 Thr Ser Ala Leu Thr Lys Lys Ile Thr Leu Lys Thr Pro Leu Val Ser
54 55 60 65
56 tct ccc atg gac aca gtc aca gag gct ggg atg gcc ata gca atg gcg 296
57 Ser Pro Met Asp Thr Val Thr Glu Ala Gly Met Ala Ile Ala Met Ala
58 70 75 80
60 ctt aca ggc ggt att ggc ttc atc cac cac aac tgt aca cct gaa ttc 344

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/846,637C

DATE: 12/26/2002

TIME: 13:32:53

Input Set : A:\2502seq.003
 Output Set: N:\CRF4\12262002\I846637C.raw

61	Leu	Thr	Gly	Gly	Ile	Gly	Phe	Ile	His	His	Asn	Cys	Thr	Pro	Glu	Phe			
62	85				90							95							
64	cag	gcc	aat	gaa	gtt	cg	aaa	gt	aag	aaa	tat	gaa	cag	gga	ttc	atc	392		
65	Gln	Ala	Asn	Glu	Val	Arg	Lys	Val	Lys	Tyr	Glu	Gln	Gly	Phe	Ile				
66	100				105				110		110				115				
68	aca	gac	cct	gt	gt	tc	tc	ag	cc	aag	gat	cgc	gt	cg	gat	ttt	440		
69	Thr	Asp	Pro	Val	Val	Leu	Ser	Pro	Lys	Asp	Arg	Val	Arg	Asp	Val	Phe			
70					120				125		125				130				
72	gag	gcc	aag	gcc	cg	cat	gg	tt	tc	tg	gt	atc	cc	atc	ac	ac	488		
73	Glu	Ala	Lys	Ala	Arg	His	Gly	Phe	Cys	Gly	Ile	Pro	Ile	Thr	Asp	Thr			
74					135				140		140				145				
76	ggc	cg	at	gg	gg	ag	cg	tt	gt	gg	gc	atc	atc	tcc	tcc	agg	gac	att	536
77	Gly	Arg	Met	Gly	Ser	Arg	Leu	Val	Gly	Ile	Ile	Ser	Ser	Arg	Asp	Ile			
78					150				155		155				160				
80	gat	ttt	ctc	aa	gag	gag	gaa	cat	gac	tgt	tt	tg	gaa	gag	ata	at	584		
81	Asp	Phe	Leu	Lys	Glu	Glu	His	Asp	Cys	Phe	Leu	Glu	Glu	Ile	Met				
82					165				170		170				175				
84	aca	aag	agg	gaa	gac	tt	gt	gt	gt	g	cc	cg	ag	atc	ac	ct	aag	632	
85	Thr	Lys	Arg	Glu	Asp	Leu	Val	Val	Ala	Pro	Arg	Ser	Ile	Thr	Leu	Lys			
86					180				185		190				195				
88	gag	gca	aat	gaa	att	ct	tg	ca	cg	ag	aag	gg	aag	tt	cc	att	680		
89	Glu	Ala	Asn	Glu	Ile	Leu	Gln	Arg	Ser	Lys	Lys	Gly	Lys	Lys	Leu	Pro	Ile		
90					200				205		205				210				
92	gta	aat	gaa	gat	gat	gag	ct	tt	gt	g	cc	atc	att	g	cc	ac	ct	728	
93	Val	Asn	Glu	Asp	Asp	Glu	Leu	Val	Ala	Ile	Ile	Ala	Arg	Thr	Asp	Leu			
94					215				220		220				225				
96	aag	aag	aat	cg	gac	ta	cc	ct	gg	cc	tcc	aaa	gat	gg	aag	aaa	cg	776	
97	Lys	Lys	Asn	Arg	Asp	Tyr	Pro	Leu	Ala	Ser	Lys	Asp	Ala	Lys	Lys	Gln			
98					230				235		235				240				
100	ct	ct	tg	gg	g	ca	gg	cc	at	gg	act	cat	gag	gat	gac	aag	tat	agg	824
101	Leu	Leu	Cys	Gly	Ala	Ala	Ile	Gly	Thr	His	Glu	Asp	Asp	Lys	Tyr	Arg			
102					245				250		250				255				
104	ct	gac	tt	ct	gc	ca	g	ct	gg	gt	gt	gt	tt	tg	gac	tct	872		
105	Leu	Asp	Leu	Leu	Ala	Gln	Ala	Gly	Val	Asp	Val	Val	Val	Leu	Asp	Ser			
106					260				265		270				275				
108	tcc	cag	gga	aat	tcc	atc	ttc	ca	tg	atc	aat	atg	atc	aag	ta	atc	aaa	920	
109	Ser	Gln	Gly	Asn	Ser	Ile	Phe	Gln	Ile	Asn	Met	Ile	Lys	Tyr	Ile	Lys			
110					280				285		285				290				
112	gac	aaa	ta	c	c	at	ct	ca	gt	gg	aa	gt	gt	gt	act	gct	968		
113	Asp	Lys	Tyr	Pro	Asn	Leu	Gln	Ile	Gly	Gly	Asn	Val	Val	Thr	Ala				
114					295				300		300				305				
116	gc	cag	gc	aa	ac	ct	cc	at	gt	gt	gt	gat	gc	ct	cg	gt	1016		
117	Ala	Gln	Ala	Lys	Asn	Leu	Ile	Asp	Ala	Gly	Val	Asp	Ala	Leu	Arg	Val			
118					310				315		315				320				
120	ggc	at	g	g	at	gg	tc	atc	tg	at	ac	ca	g	gt	ct	gg	tgt	1064	
121	Gly	Met	Gly	Ser	Gly	Ser	Ile	Cys	Ile	Thr	Gln	Glu	Val	Leu	Ala	Cys			
122					325				330		330				335				
124	gg	cg	cc	ca	ca	ca	ca	gt	tg	1112									
125	Gly	Arg	Pro	Gln	Ala	Thr	Ala	Val	Tyr	Lys	Val	Ser	Glu	Tyr	Ala	Arg			

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/846,637C

DATE: 12/26/2002
 TIME: 13:32:53

Input Set : A:\2502seq.003
 Output Set: N:\CRF4\12262002\I846637C.raw

126	340	345	350	355	
128	cgc ttt ggt gtt ccg gtc att gct gat gga gga atc caa aat gtg ggt				1160
129	Arg Phe Gly Val Pro Val Ile Ala Asp Gly Gly Ile Gln Asn Val Gly				
130	360	365	370		
132	cat att gcg aaa gcc ttg gcc ctt ggg gcc tcc aca gtc atg atg ggc				1208
133	His Ile Ala Lys Ala Leu Ala Leu Gly Ala Ser Thr Val Met Met Gly				
134	375	380	385		
136	tct ctc ctg gct gcc acc act gag gcc cct ggt gaa tac ttc ttt tcc				1256
137	Ser Leu Leu Ala Ala Thr Thr Glu Ala Pro Gly Glu Tyr Phe Phe Ser				
138	390	395	400		
140	gat ggg atc cgg cta aag aaa tat cgc ggt atg ggt tct ctc gat gcc				1304
141	Asp Gly Ile Arg Leu Lys Tyr Arg Gly Met Gly Ser Leu Asp Ala				
142	405	410	415		
144	atg gac aag cac ctc agc agc aac aga tat ttc agt gaa gct gac				1352
145	Met Asp Lys His Leu Ser Ser Gln Asn Arg Tyr Phe Ser Glu Ala Asp				
146	420	425	430	435	
148	aaa atc aaa gtg gcc cag gga gtg tct ggt gct gtg cag gac aaa ggg				1400
149	Lys Ile Lys Val Ala Gln Gly Val Ser Gly Ala Val Gln Asp Lys Gly				
150	440	445	450		
152	tca atc cac aaa ttt gtc cct tac ctg att gct ggc atc caa cac tca				1448
153	Ser Ile His Lys Phe Val Pro Tyr Leu Ile Ala Gly Ile Gln His Ser				
154	455	460	465		
156	tgc cag gac att ggt gcc aag agc ttg acc caa gtc cga gcc atg atg				1496
157	Cys Gln Asp Ile Gly Ala Lys Ser Leu Thr Gln Val Arg Ala Met Met				
158	470	475	480		
160	tac tct ggg gag ctt aag ttt gag aag aga acg tcc tca gcc cag gtg				1544
161	Tyr Ser Gly Glu Leu Lys Phe Glu Lys Arg Thr Ser Ser Ala Gln Val				
162	485	490	495		
164	gaa ggt ggc gtc cat agc ctc cat tcg tat gag aag cgg ctt ttc				1589
165	Glu Gly Gly Val His Ser Leu His Ser Tyr Glu Lys Arg Leu Phe				
166	500	505	510		
168	tgaaaaggga tccagcacac ctcctcggtt ttttttcaa taaaagttta gaaagaccg				1649
169	aattc				1654
171	<210> SEQ ID NO: 2				
172	<211> LENGTH: 514				
173	<212> TYPE: PRT				
174	<213> ORGANISM: Homo sapien				
176	<400> SEQUENCE: 2				
177	Met Ala Asp Tyr Leu Ile Ser Gly Gly Thr Ser Tyr Val Pro Asp Asp				
178	1	5	10	15	
179	Gly Leu Thr Ala Gln Gln Leu Phe Asn Cys Gly Asp Gly Leu Thr Tyr				
180	20	25	30		
181	Asn Asp Phe Leu Ile Leu Pro Gly Tyr Ile Asp Phe Thr Ala Asp Gln				
182	35	40	45		
183	Val Asp Leu Thr Ser Ala Leu Thr Lys Lys Ile Thr Leu Lys Thr Pro				
184	50	55	60		
185	Leu Val Ser Ser Pro Met Asp Thr Val Thr Glu Ala Gly Met Ala Ile				
186	65	70	75	80	
187	Ala Met Ala Leu Thr Gly Gly Ile Gly Phe Ile His His Asn Cys Thr				

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/846,637C

DATE: 12/26/2002

TIME: 13:32:53

Input Set : A:\2502seq.003

Output Set: N:\CRF4\12262002\I846637C.raw

188	85	90	95	
189	Pro Glu Phe Gln Ala Asn Glu Val Arg Lys Val Lys Lys Tyr Glu Gln			
190	100	105	110	
191	Gly Phe Ile Thr Asp Pro Val Val Leu Ser Pro Lys Asp Arg Val Arg			
192	115	120	125	
193	Asp Val Phe Glu Ala Lys Ala Arg His Gly Phe Cys Gly Ile Pro Ile			
194	130	135	140	
195	Thr Asp Thr Gly Arg Met Gly Ser Arg Leu Val Gly Ile Ile Ser Ser			
196	145	150	155	160
197	Arg Asp Ile Asp Phe Leu Lys Glu Glu Glu His Asp Cys Phe Leu Glu			
198	165	170	175	
199	Glu Ile Met Thr Lys Arg Glu Asp Leu Val Val Ala Pro Arg Ser Ile			
200	180	185	190	
201	Thr Leu Lys Glu Ala Asn Glu Ile Leu Gln Arg Ser Lys Lys Gly Lys			
202	195	200	205	
203	Leu Pro Ile Val Asn Glu Asp Asp Glu Leu Val Ala Ile Ile Ala Arg			
204	210	215	220	
205	Thr Asp Leu Lys Lys Asn Arg Asp Tyr Pro Leu Ala Ser Lys Asp Ala			
206	225	230	235	240
207	Lys Lys Gln Leu Leu Cys Gly Ala Ala Ile Gly Thr His Glu Asp Asp			
208	245	250	255	
209	Lys Tyr Arg Leu Asp Leu Leu Ala Gln Ala Gly Val Asp Val Val Val			
210	260	265	270	
211	Leu Asp Ser Ser Gln Gly Asn Ser Ile Phe Gln Ile Asn Met Ile Lys			
212	275	280	285	
213	Tyr Ile Lys Asp Lys Tyr Pro Asn Leu Gln Val Ile Gly Gly Asn Val			
214	290	295	300	
215	Val Thr Ala Ala Gln Ala Lys Asn Leu Ile Asp Ala Gly Val Asp Ala			
216	305	310	315	320
217	Leu Arg Val Gly Met Gly Ser Gly Ser Ile Cys Ile Thr Gln Glu Val			
218	325	330	335	
219	Leu Ala Cys Gly Arg Pro Gln Ala Thr Ala Val Tyr Lys Val Ser Glu			
220	340	345	350	
221	Tyr Ala Arg Arg Phe Gly Val Pro Val Ile Ala Asp Gly Gly Ile Gln			
222	355	360	365	
223	Asn Val Gly His Ile Ala Lys Ala Leu Ala Leu Gly Ala Ser Thr Val			
224	370	375	380	
225	Met Met Gly Ser Leu Leu Ala Ala Thr Thr Glu Ala Pro Gly Glu Tyr			
226	385	390	395	400
227	Phe Phe Ser Asp Gly Ile Arg Leu Lys Lys Tyr Arg Gly Met Gly Ser			
228	405	410	415	
229	Leu Asp Ala Met Asp Lys His Leu Ser Ser Gln Asn Arg Tyr Phe Ser			
230	420	425	430	
231	Glu Ala Asp Lys Ile Lys Val Ala Gln Gly Val Ser Gly Ala Val Gln			
232	435	440	445	
233	Asp Lys Gly Ser Ile His Lys Phe Val Pro Tyr Leu Ile Ala Gly Ile			
234	450	455	460	
235	Gln His Ser Cys Gln Asp Ile Gly Ala Lys Ser Leu Thr Gln Val Arg			
236	465	470	475	480

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/846,637C

DATE: 12/26/2002

TIME: 13:32:53

Input Set : A:\2502seq.003
 Output Set: N:\CRF4\12262002\I846637C.raw

237 Ala Met Met Tyr Ser Gly Glu Leu Lys Phe Glu Lys Arg Thr Ser Ser
 238 485 490 495
 239 Ala Gln Val Glu Gly Gly Val His Ser Leu His Ser Tyr Glu Lys Arg
 240 500 505 510
 241 Leu Phe
 244 <210> SEQ ID NO: 3
 245 <211> LENGTH: 1654
 246 <212> TYPE: DNA
 247 <213> ORGANISM: Homo sapien
 249 <220> FEATURE:
 250 <221> NAME/KEY: CDS
 251 <222> LOCATION: (48)...(1589)
 252 <223> OTHER INFORMATION: T333I/S351Y Human IMPDH II mutant
 254 <221> NAME/KEY: mutation
 255 <222> LOCATION: 1045
 256 <223> OTHER INFORMATION: C to T mutation

W--> 258 <221> mutation
 259 <222> LOCATION: 1046
 260 <223> OTHER INFORMATION: G to C mutation

W--> 262 <221> mutation
 263 <222> LOCATION: 1099
 264 <223> OTHER INFORMATION: C to A mutation

W--> 266 <221> mutation
 267 <222> LOCATION: 1100
 268 <223> OTHER INFORMATION: A to T mutation

W--> 270 <400> 3

271	gaattcgggc ggtcctcgga gacacgcggc ggtgtcctgt gttggcc atg gcc gac	56
272	Met Ala Asp	
273	1	
275	tac ctg att agt ggg ggc acg tcc tac gtg cca gac gac gga ctc aca	104
276	Tyr Leu Ile Ser Gly Gly Thr Ser Tyr Val Pro Asp Asp Gly Leu Thr	
277	5 10 15	
279	gca cag cag ctc ttc aac tgc gga gac ggc ctc acc tac aat gac ttt	152
280	Ala Gln Gln Leu Phe Asn Cys Gly Asp Gly Leu Thr Tyr Asn Asp Phe	
281	20 25 30 35	
283	ctc att ctc cct ggg tac atc gac ttc act gca gac cag gtg gac ctg	200
284	Leu Ile Ile Pro Gly Tyr Ile Asp Phe Thr Ala Asp Gln Val Asp Leu	
285	40 45 50	
287	act tct gct ctg acc aag aaa atc act ctt aag acc cca ctg gtt tcc	248
288	Thr Ser Ala Leu Thr Lys Lys Ile Thr Leu Lys Thr Pro Leu Val Ser	
289	55 60 65	
291	tct ccc atg gac aca gtc aca gag gct ggg atg gcc ata gca atg gcg	296
292	Ser Pro Met Asp Thr Val Thr Glu Ala Gly Met Ala Ile Ala Met Ala	
293	70 75 80	
295	ctt aca ggc ggt att ggc ttc atc cac cac aac tgt aca cct gaa ttc	344
296	Leu Thr Gly Gly Ile Gly Phe Ile His His Asn Cys Thr Pro Glu Phe	
297	85 90 95	
299	cag gcc aat gaa gtt cgg aaa gtg aag aaa tat gaa cag gga ttc atc	392
300	Gln Ala Asn Glu Val Arg Lys Val Lys Lys Tyr Glu Gln Gly Phe Ile	

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/846,637C

DATE: 12/26/2002
TIME: 13:32:54

Input Set : A:\2502seq.003
Output Set: N:\CRF4\12262002\I846637C.raw

L:258 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:262 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:266 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:270 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:500 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:5
L:730 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:7
L:953 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:9
L:965 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:9
L:1188 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:11
L:1199 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:11
L:2109 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#:19, CDS LOCATION: (0)...
(1188)
L:2462 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:23
L:2473 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:23
L:2634 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:25
~~L:2855 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:27~~
L:3297 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:31
L:3301 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:31